Read Me

Because of the nature of the XML parser I was using I had to make 2 small modifications with the input files to get them in proper xml. I have provided the changes in the files and I have listed them below so you can check with your own files if you please.

1) In line-train.txt, line 1866, instance: line-n.w7_098:4032

   <instance id="line-n.w7_098:4032:">

   <answer instance="line-n.w7_098:4032:" senseid="phone"/>

   <context>

   <s> "It's a step in the right direction," an investment banker said of the move. </s> <@> </p> <@> <p> <@> <s> A Fed spokesman said Friday that extending <head>lines</head> to investment banks is still under consideration. </s> </s> <@> </p> <@> <p> <@> <s> Morgan Stanley & Co., Goldman, Sachs & Co. and Salomon Brothers Inc. are among the investment banks that have become major foreign exchange market participants in recent years and seek lines to the Fed. </s> </s> <@> </p> <@> <p> <@> <s> The Fed wouldn't say which or how many brokers it established links with, but two New York brokers identified themselves as having opened lines to the Fed: Noonan, Astley & Pearce Inc. and Tullett & Tokyo Forex Inc. </s> <@> </p> <@> <p> <@> <s> For the Fed, lines to brokers provide instantaneous information on exchange-rate moves. </s> <@> <s> For brokers, as for commercial banks, getting a line to the Fed brings prestige and at least the appearance of having access to the internal workings of the often secretive central bank. </s> <@> </p> <@> <p> <@> <s> If the Fed eventually chooses to use brokers for its currency transactions, the lines also could deliver sizable chunks of profitable business. </s>

   </context>

   </instance>

   This is instance for some reason is to only instance in the entire training data that is not formatted correctly. I tried many of ways to automatically format it, but the easiest was **just to remove it.**

2) The only change with the key data is just to **add start and end tags**. I just added <answers> and </answers> to the end of the file.